Electron fake rates and trigger efficiency

Mike Flowerdew

michael.flowerdew@cern.ch

University of Liverpool

Measuring trigger efficiencies from data: Introduction

- *The Aim:* To reduce dependency on MC, and measure (relative) trigger efficiencies from real data
- Method known as “tag and probe” or “double object” method
- What is measured? The *single* object trigger efficiency, relative to the offline selection
- For e25i, we use $Z \rightarrow ee$ decays
- This work has been in collaboration with many people from CERN, including E. Dobson, N. Ellis, T. Fonseca Martin, C. Padilla and T. Weidberg, and my supervisor at Liverpool, Joost Vossebeld

The Art of Statistical Computing

- There is a large amount of available data available on the grid for sample 5144 (Pythia, inclusive $Z \to ee$)
  - `trig1_misal1_csc11.005144.PythiaZee.recon.AOD.v12000601`
  - Sample includes material distortions and misaligned geometry
  - In v12, full trigger information is on ESD and AOD
- AODs used to produce custom NTuple, with full electron trigger slice included
- DQ2 input dataset $\Rightarrow$ use ganga
- Dataset split into 100 subjobs:
  - 17 - Job proxy expired
  - 52 - Error copying output file
  - 31 - Successful
  - Total output retrieved - 25
  - 105,000 events

The tag and probe method

In its simplest form, this method integrates over all relevant kinematical variables to obtain a *global* trigger efficiency. Two samples are defined:

- *Diagnostic sample:* N_1 events where at least one electron passes trigger
- *Control sample:* N_2 events where at least two electrons pass trigger

The numbers N_1 and N_2 are determined by counting in the absence of background (more generally, by sideband subtraction or a mass peak fit).

To give a clean enough sample, insist on offline event selection with efficiency ϵ_r. Then, given N_0 true events, with acceptance $\mathcal{A}$:

$$
\begin{aligned}
N_1 - B_1 &= (2\epsilon_t - \epsilon_t^2) \cdot \epsilon_r \mathcal{A} N_0 \\
N_2 - B_2 &= \epsilon_t^2 \cdot \epsilon_r \mathcal{A} N_0 \\
\epsilon_t &= \frac{2(N_2 - B_2)}{(N_1 - B_1) + (N_2 - B_2)}
\end{aligned}
$$

Global trigger efficiencies

Offline selection:

- 2 loose electrons (cluster + track match + $E_T > 25$ GeV)
- Opposite charge
- $70 < M_{ee} < 100$ GeV (Radiative losses & EM energy scale - G. Unal)

All efficiencies are relative to the offline selection and double-object signature of all lower levels

Trigger Level	With crack	Without crack	Preselection
L1	(97.6±0.1)%	(97.7±0.1)%	Offline + EM25I
L2	(94.0±0.1)%	(95.7±0.1)%	Offline + 2EM25I + e25i(L2)
EF	(94.3±0.1)%	(94.4±0.1)%	Offline + 2EM25I + 2e25i(L2) + e25i(EF)

Differential trigger efficiency

The global efficiency is essentially a differential efficiency with one bin. With more than one bin, different statistics apply depending on where the electrons fall. Consider probe electrons in *one* bin:

- Remember, N_1 and N_2 count *events*, not *electrons*
- Case A: Tag is in a different bin - electrons are distinguishable

$$\epsilon \sim N_2/N_1$$

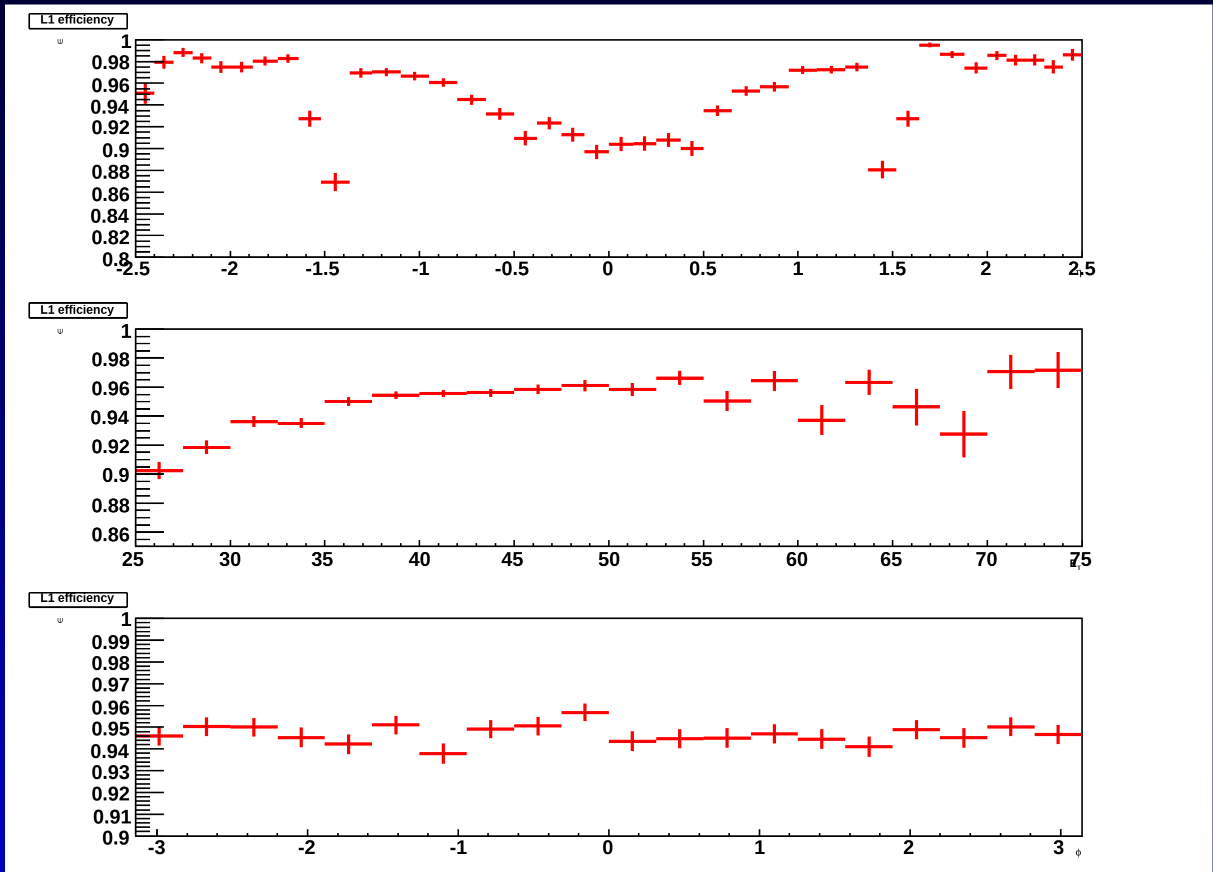
- Case B: Tag is in the same bin - indistinguishable, like the global case

$$\epsilon \sim 2N_2/(N_1 + N_2)$$

$$\begin{aligned}
\epsilon &= \frac{N_2^A + 2N_2^B}{N_1^A + N_1^B + N_2^B} = \frac{N_2^A + 2N_2^B}{N_T}, \\
\sigma_\epsilon &= \sqrt{\frac{(1-\epsilon)}{N_T}\left[\epsilon + \frac{2N_2^B}{N_T}(1-\epsilon)\right]}
\end{aligned}$$

Equation depends on binning, not how many dimensions the histogram has

L1 trigger efficiency (EM25I)



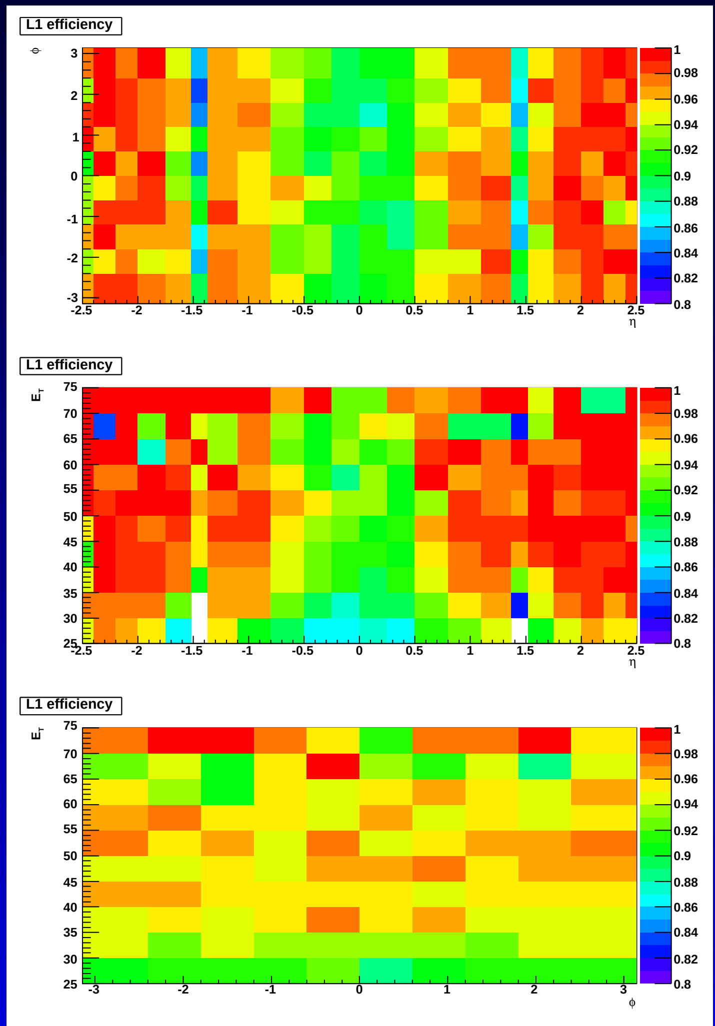
η

E_T

ϕ

Dip at low η is caused by isolation cuts (see T. Fonseca Martin's egamma talk from Tuesday)

L1 trigger efficiency (EM25I)



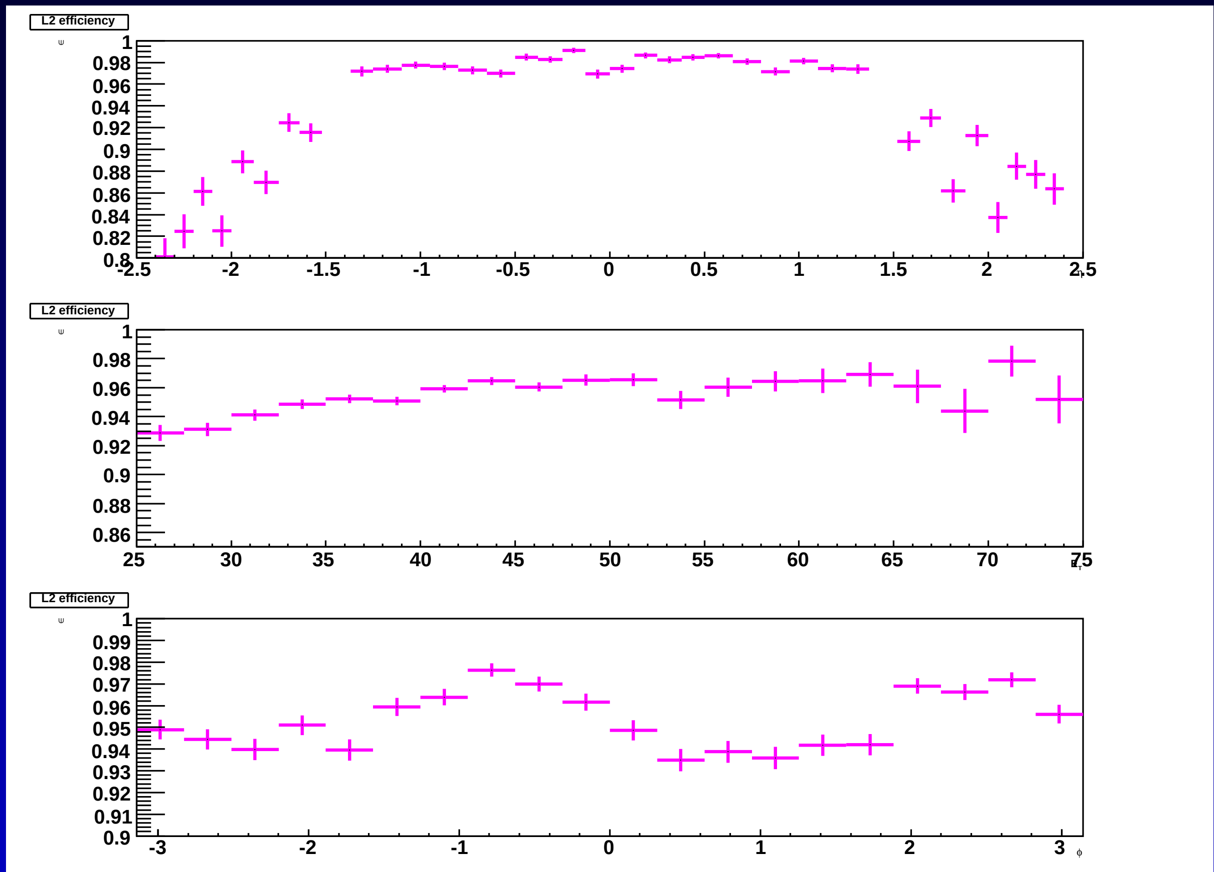
η vs. ϕ

η vs. p_{T}

ϕ vs. p_{T}

Error at 90% efficiency is $\sim 1 - 2\%$

L2 trigger efficiency (e25i)



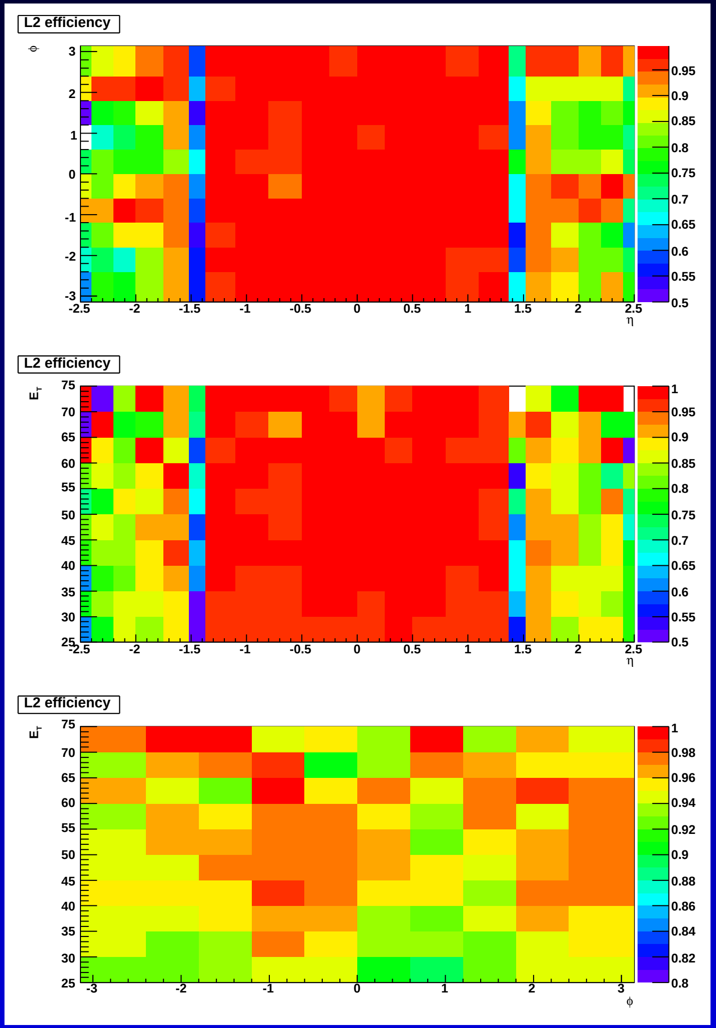
η

E_T

ϕ

2D plots may shed some light on these strange shapes…

L2 trigger efficiency (EM25I)



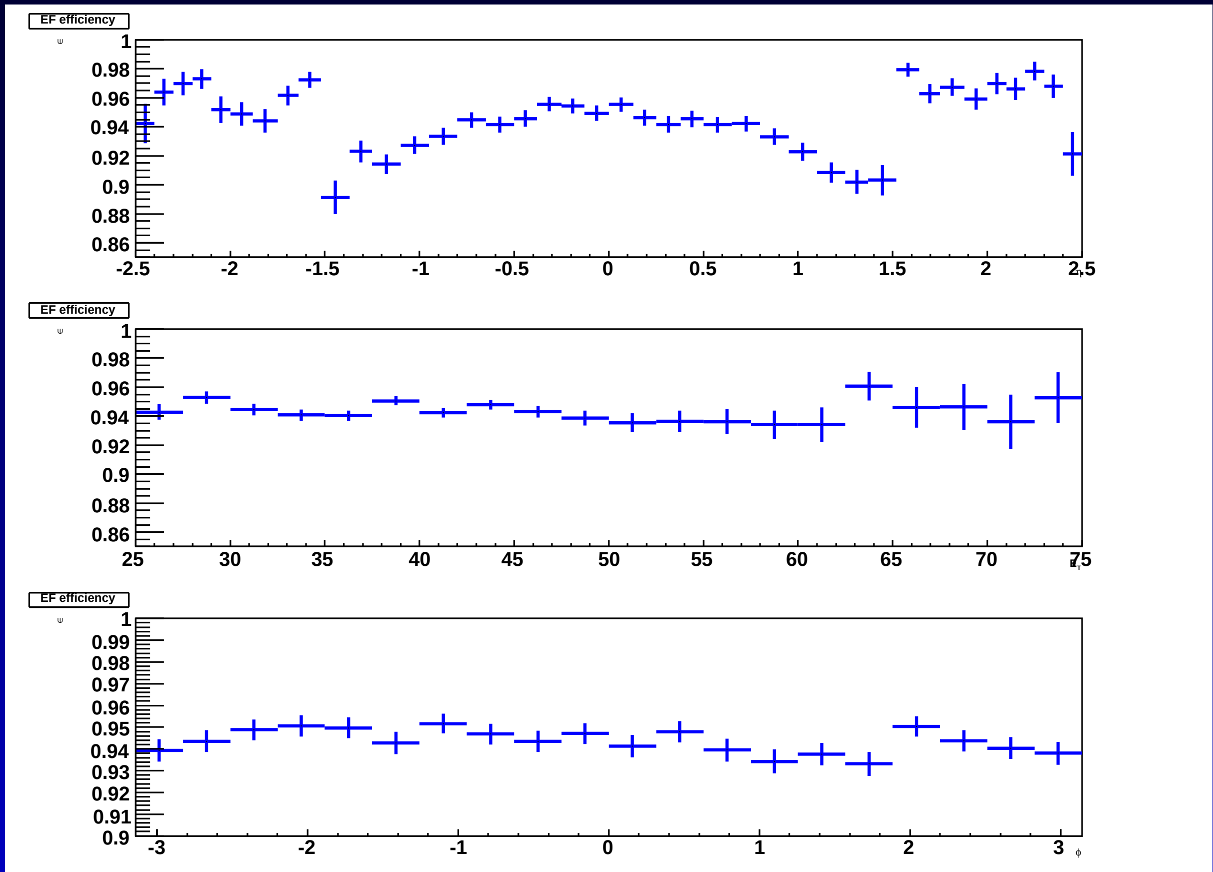
η vs. ϕ

η vs. p_{T}

ϕ vs. p_{T}

Error at 90% efficiency is $\sim$ 2%

EF trigger efficiency (e25i)



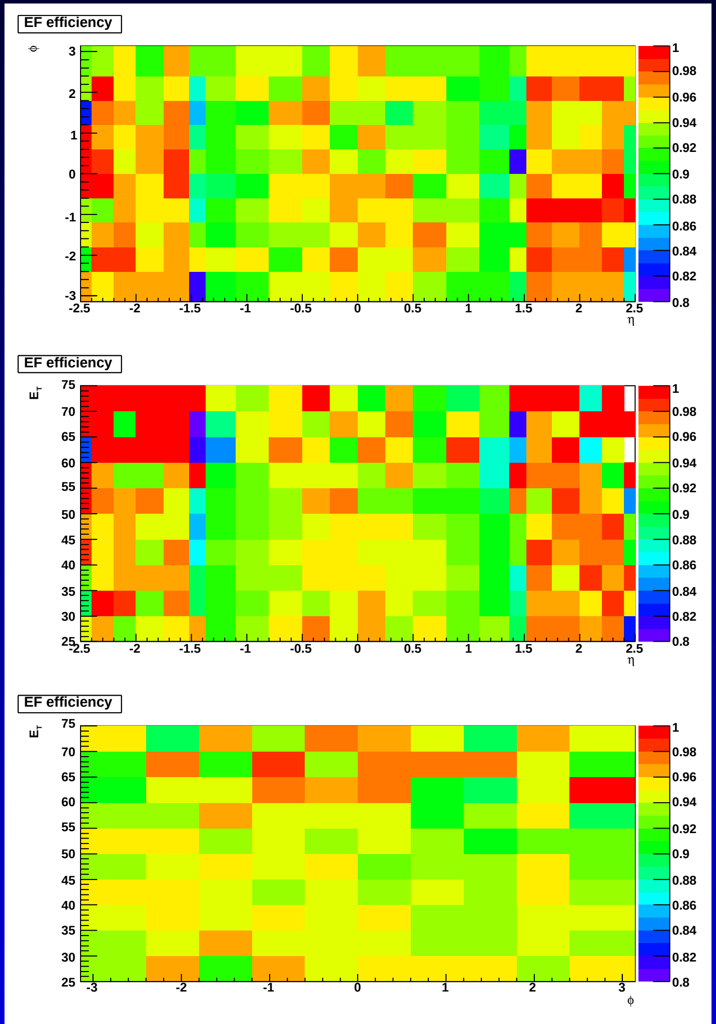
η

E_T

ϕ

Efficiency is flat in p_{T}, but not in η. But this is relative to L2, which was very efficient in the barrel, and not in the end-caps

EF trigger efficiency (EM25I)



η vs. ϕ

η vs. p_{T}

ϕ vs. p_{T}

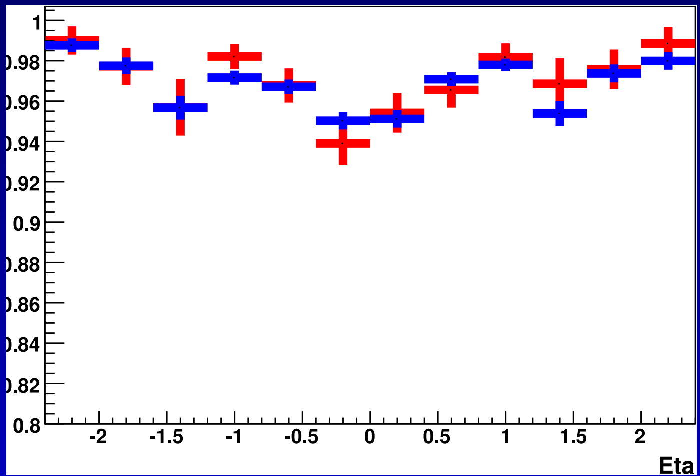
Error at 90% efficiency is $\sim$ 2%

Further work: Trigger efficiency measurement

Starting to obtain useful statistics for differential trigger studies.
Varied interests within the group:

Ellie Z+jets, effect of jet p_{T}, electron/jet spatial separation, jets in cracks...

Teresa Systematics, eg using a different calibration for HLT and offline reconstruction



Comparison of misaligned geometry vs. ideal (T. Fonseca Martin)
Red = Ideal
Blue = Misaligned
Also no noticeable variation with definition of offline electron (loose/tight)

Tony Offline reconstruction efficiency, starting with e-jet sample

Myself NTuple rest of $Z \rightarrow ee$ dataset - $5\times$ what I have now, and investigate features in differential plots and backgrounds

Estimating the jet background in $Z \rightarrow ee$: Introduction

This study uses version 11.0.42
The aim: to effectively parameterise backgrounds to the $Z \rightarrow ee$ channel involving QCD fakes.
Here a fake is anything that isn't a real electron from an electroweak process
The idea:

1. Parameterise the probability of a single jet to be reconstructed as an electron (the *Jet Weight*)
2. Use this parameterisation to produce an *Event Weight* - a product of two Jet Weights in the case of dijet events

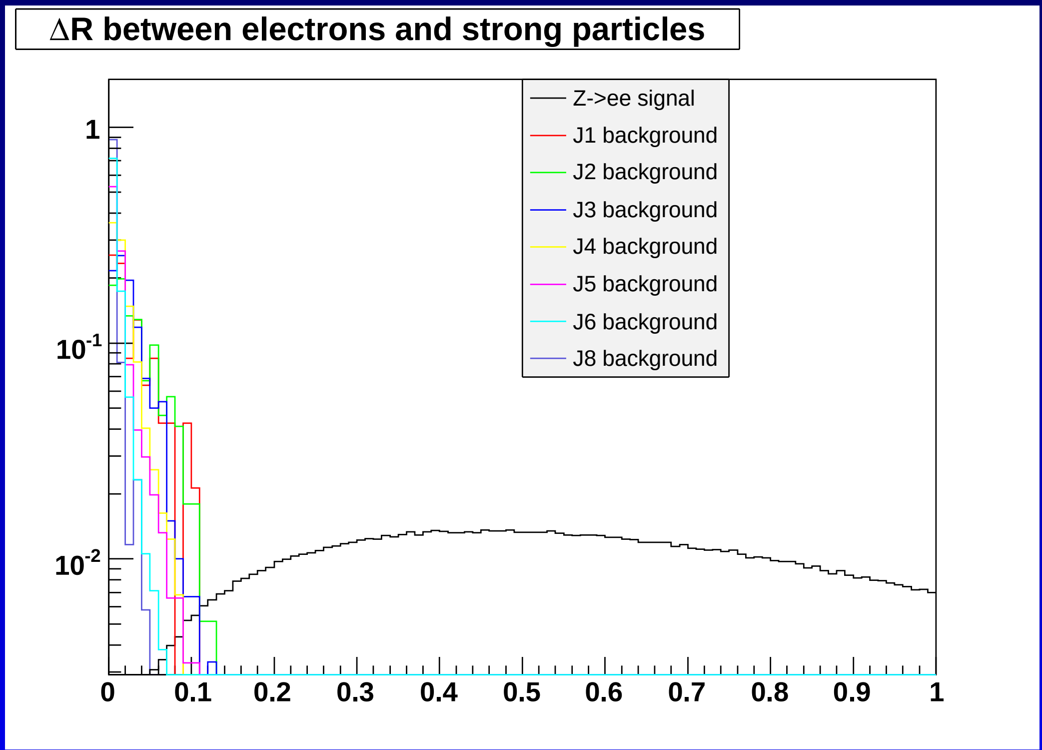
Samples ntupled via the grid:

- Pythia $Z \rightarrow ee$ sample 5144 - 490,000 events
- Pythia jets (J1-J8), samples 5010-5017 (J7 missing) - 258,000 events in J2
- Pythia $W \rightarrow e\nu$, sample 5104 - 90,000 events

Official AOD production, no trigger information

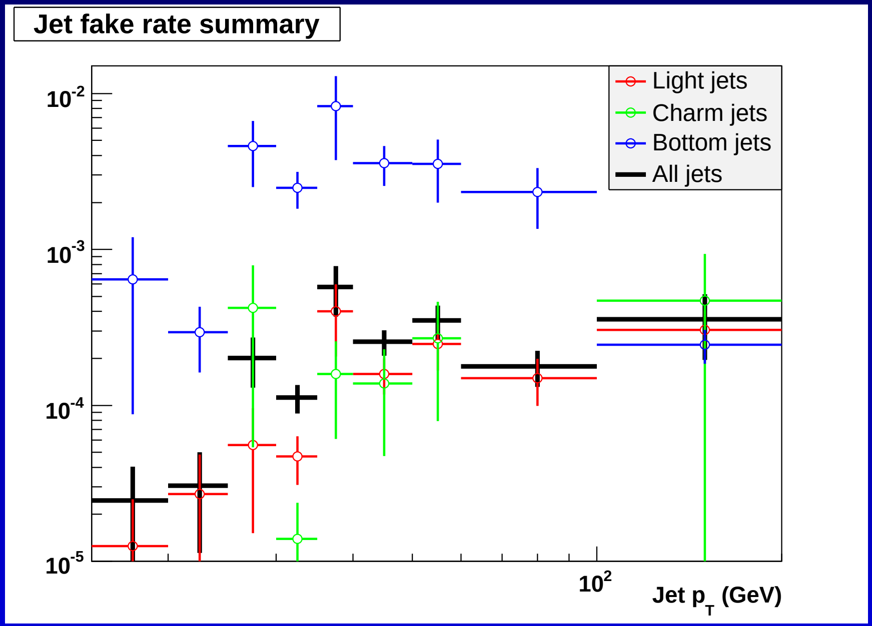
Determining the Jet Weight: Numerator

Efficiency is a ratio: N_{elec}/N_{jets}.
Plot shows separation (ΔR) between true electron and nearest QCD particle.
Real electrons in dijet events are associated with heavy flavour decays and will be close to jets.
Numerator: Fiducial reconstructed electron with isEM==0, $E_T > 25$ GeV, $\Delta\mathrm{R}(\mathrm{elec}, \mathrm{jet}) < 0.4$.

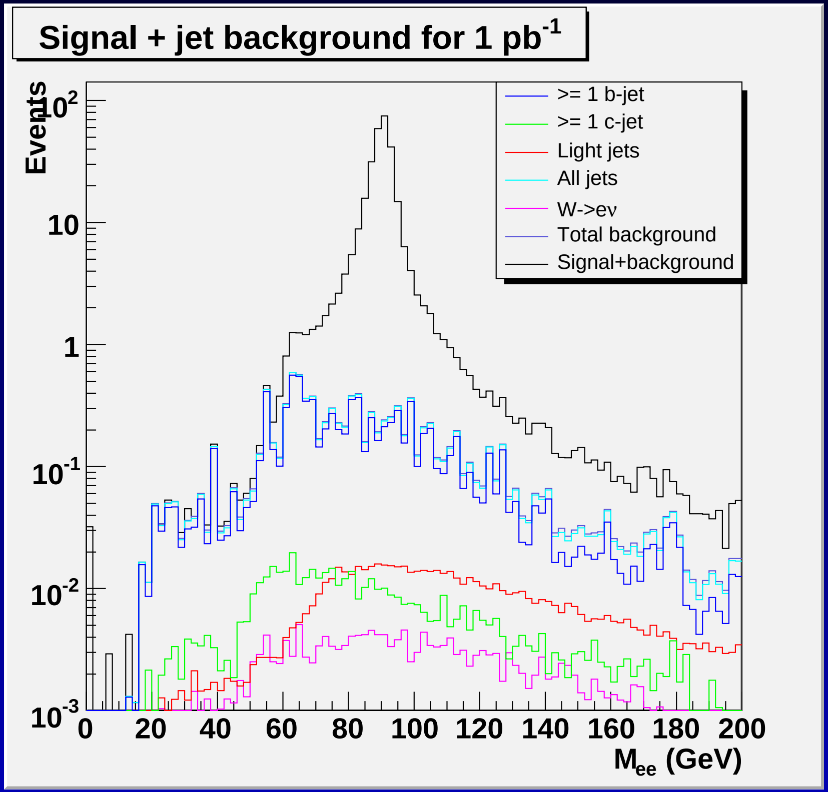


Determining the Jet Weight: Denominator and result

Efficiency is a ratio: N_{elec}/N_{jets}.
Heavy flavour jets may contain real electrons $\Rightarrow$ these need to be dealt with separately - JetTagInfo associates each jet with the truth.
BJetCollection is essentially a copy of Cone4TowerParticleJets, with extra flavour information.
Denominator: BJetCollection jet, separated by jet type and p_{T}.



Background to the $Z \rightarrow ee$ channel



For dijets:
Event Weight =
JetWeight1(p_{T}, flavour)
$\times$ JetWeight(p_{T}, flavour)

For $W \rightarrow e\nu$, e-jet combinations are used:
Event Weight =
JetWeight(p_{T},flavour)

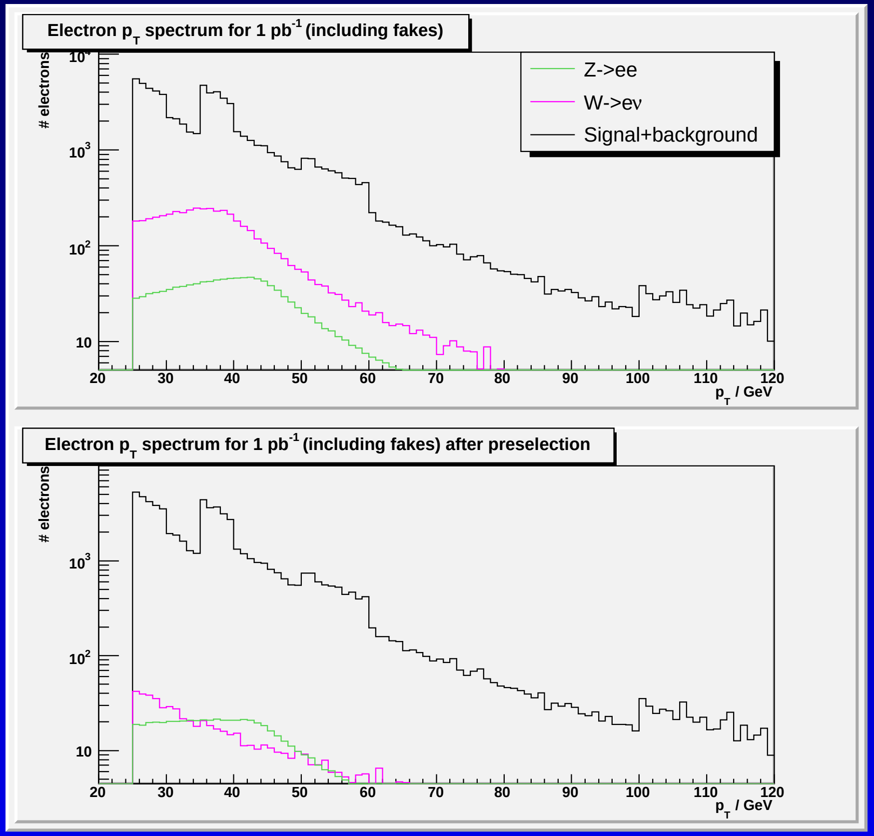
Health Warning: No allowance made here for difference between jet and EM energy scales.
Luminosity = $1\mathrm{pb}^{-1}$, with 284 signal events, 7.6 background events.
($61 < M_Z < 121$ GeV)

Can Jet Weights be measured from data?

Must find an unbiased jet sample which is sufficiently pure.

Top plot: Reconstructed single electron E_T spectrum (using Jet Weights)

Bottom plot: Anti-W/Z cuts applied: $E_T^{miss} < 25$ GeV and
< 2 reconstructed electrons



Further work: Jet background in $Z \longrightarrow ee$

- Version 12.0.6 AODs now available, new data appearing all the time
  - What effect will distorted geometry have on these results?
- Grid ntupling is very inefficient (see earlier) - getting meaningful statistics may take time as jobs fail and are resubmitted...
- Available statistics with filtered dijets (5802) may be more significant than Pythia Jn samples used so far
- Many effects still to study:
  - Include effect of trigger
  - Isolation to improve jet rejection (stats?)
  - Hard (leading) jet vs. softer (sub-leading) jets
  - Effect of different cone sizes
  - Backgrounds for loose/medium/tight offline electron selections